

>human SULF1 full length cDNA (ORF highlighted in capitals)

ccaccaccatcatctaaagaagataaacttggcgaatgacatgcaggttcttcaaggcagaataattgcagaaaatcttcaaa
ggaccctatctgcagatgttctgaatacctctgagaatagagattgattatcaaccaggatacctaattcaaggactccagaaat
caggagacggagacatttgtcagtttgcacattggaccaaaaatacaATGAAGTATTCTTGCTGTGCTCTGG
TTTTGGCTGTCCTGGGCACAGAATTGCTGGGAAGCCTCTGTTGCGACTGTCAGATCCCC
GAGGTTTCAGAGGACGGATACAGCAGGAACGAAAAACATCCGACCCAACATTATTCTTG
TGCTTACCGATGATCAAGATGTGGAGCTGGGGTCCCTGCAAGTCATGAACAAAACGAG
AAAGATTATGGAACATGGGGGGGCCACCTTCATCAATGCCTTTGTGACTACACCCATGT
GCTGCCCCGTCACGGTCCTCCATGCTCACCGGGAAGTATGTGCACAATCACAATGTCTA
CACCACAACGAGAAGTCTCTTCCCCCTCGTGGCAGGCCATGCATGAGCCTCGGACT
TTTGCTGTATATCTTAACAACACTGGCTACAGAACAGCCTTTTTTGGAAAATACCTCAATG
AATATAATGGCAGCTACATCCCCCTGGGTGGCGAGAATGGCTTGGATTAATCAAGAATT
CTCGCTTCTATAATTACACTGTTTGTGCGCAATGGCATCAAAGAAAAGCATGGATTGATTA
TGCAAAGGACTACTTCACAGACTTAATCACTAACGAGAGCATTAACTTCAAAATGTCT
AAGAGAATGTATCCCCATAGGCCCGTTATGATGGTGATCAGCCACGCTGCGCCCCACG
GCCCCGAGGACTCAGCCCCACAGTTTTCTAAACTGTACCCCAATGCTTCCCAACACATA
ACTCCTAGTTATAACTATGCACCAAATATGGATAAACACTGGATTATGCAGTACACAGGAC
CAATGCTGCCCATCCACATGGAATTTACAAACATTCTACAGCGCAAAAGGCTCCAGACT
TTGATGTCAGTGGATGATTCTGTGGAGAGGCTGTATAACATGCTCGTGGAGACGGGGG
AGCTGGAGAATACTTACATCATTTACACCGCCGACCATGGTTACCATATTGGGCAGTTTG
GACTGGTCAAGGGGAAATCCATGCCATATGACTTTGATATTGCTGTGCCTTTTTTTTATTC
GTGGTCCAAGTGTAGAACAGGATCAATAGTCCCACAGATCGTTCTCAACATTGACTTG
GCCCCACGATCCTGGATATTGCTGGGCTCGACACACCTCCTGATGTGGACGGCAAGT
CTGTCTCAAACCTTCTGGACCCAGAAAAGCCAGGTAACAGGTTTCGAACAAACAAGAA
GGCCAAAATTTGGCGTGATACATTCTAGTGGAAGAGGCAAATTTCTACGTAAGAAGG
AAGAATCCAGCAAGAATATCCAACAGTCAAATCACTTGCCCAAATATGAACGGGTCAA
GAACTATGCCAGCAGGCCAGGTACCAGACAGCCTGTGAACAACCGGGGCAGAAGTGG
CAATGCATTGAGGATACATCTGGCAAGCTTCGAATTCACAAGTGTAAAGGACCCAGTGA
CCTGCTCACAGTCCGGCAGAGCACGCGGAACCTCTACGCTCGCGGCTTCCATGACAA
AGACAAAGAGTGCAGTTGTAGGGAGTCTGGTTACCGTGCCAGCAGAAGCCAAAGAAAG
AGTCAACGGCAATTCTTGAGAAACCAGGGGACTCCAAAGTACAAGCCCAGATTTGTCC
ATACTCGGCAGACACGTTCTTGTCCGTCGAATTTGAAGGTGAAATATATGACATAAATC
TGGAAGAAGAAGAAGATTGCAAGTGTTGCAACCAAGAAACATTGCTAAGCGTCATGAT
GAAGGCCACAAGGGGCCAAGAGATCTCCAGGCTTCCAGTGGTGGCAACAGGGGCAG
GATGCTGGCAGATAGCAGCAACGCCGTGGGCCACCTACCACTGTCCGAGTGACACA
CAAGTGTTTTATTCTTCCCAATGACTCTATCCATTGTGAGAGAGAAGTGTACCAATCGGC
CAGAGCGTGGAAGGACCATAAGGCATACATTGACAAAGAGATTGAAGCTCTGCAAGATA
AAATTAAGAATTTAAGAGAAGTGAGAGGACATCTGAAGAGAAGGAAGCCTGAGGAATGT
AGCTGCAGTAAACAAAGCTATTACAATAAAGAGAAAGGTGTAAAAAGCAAGAGAAATTA
AAGAGCCATCTTCACCCATTCAAGGAGGCTGCTCAGGAAGTAGATAGCAAACCTGCAACT
TTTCAAGGAGAACAACCGTAGGAGGAAGAAGGAGAGGAAGGAGAAGAGACGGCAGAG
GAAGGGGGAAGAGTGCAGCCTGCCTGGCCTCACTTGCTTCACGCATGACAACAACCA
CTGGCAGACAGCCCCGTTCTGGAACCTGGGATCTTTCTGTGCTTGCACGAGTTCTAAC
AATAACACCTACTGGTGTTTTCGCTACAGTTAATGAGACGCATAATTTTCTTTTCTGTGAGT
TTGCTACTGGCTTTTTGGAGTATTTTGAATACAGATCCTTATCAGCTCACAAATAC
AGTGACACCGGTAGAACGAGGCATTTTGAATCAGCTACACGTACAATAATGGAGCTCA
GAAGCTGTCAAGGATATAAGCAGTGCAACCCAAGACCTAAGAATCTTGATGTTGGAAAT
AAAGATGGAGGAAGCTATGACCTACACAGAGGACAGTTATGGGATGGATGGGAAGGTT

FIG. 1Ai

AAAtcagccccgtctcactgcagacatcaactggcaaggcctagaggagctacacagtgtgaatgaaaacat
ctatgagtacagacaaaactacagacttagtctgggtggactggactaattacttgaaggatttagatagagtattt
gcactgctgaagagtcactatgagcaaaaataaaacaaataagactcaaaactgctcaaagtgcgggttcttg
gtgtgtctgctgagcacgctgtgtcaatggagatggcctctgctgactcagatgaagaccaaggcataaggt
tgggaaaacacctcatttgacctgacgtgaccttcaaaccctgcatttgaaccgaccaacattaagtccag
agagttaaacttgaatggaataacgacatccagaagttaatcatttgaattctgaacactggagaaaaaccga
aaaatggacggggcatgaagagactaatcatctggaaaccgatttcagtggcgatggcatgacagagctag
agctcggggcccagccccaggctgcagccattcgaggcaccgaaagaacttcccagtatgggtgctct
ggaaaggacattttgaagatcaactatcttctgtgcatccgatggaatttcagttcatcagatgttcaccatg
gccaccgcagaacaccgaagtaattccagcatagcggggaagatgttgaccaagggtggagaagaatcac
gaaaaggagaagtcacagcacctagaaggcagcgctcctctcactctcctctgattagatgaaactgttac
cttaccctaaacacagatatttctttaactttttatttgaactaataaaggtaatcacagccaccaacattccaa
gctaccctgggtaccttgtgcagtagaagctagtgcagcatgtgagcaagcgggtgtgcacacgggagactcatc
gttataattactatctgccaagagtagaaagaaaggctggggatattgggtggcttgggtttgattttgtgttt
gtttgtttgtactaaaacagtattatctttgaatcgttagggacataagtatatacatgttatccaatcaagatgg
ctagaatgggtccttctgagtgtctaaaactgacacccctggtaaatcttcaacacacttccactgcctgcgta
atgaagtttgattcattttaaccactggaattttcaatgccgtcatttcagttagatgatttgcactttgagattaa
aatgccatgtctatttgattagcttattttttttttacaggcttatcagtcctcactgttggtgtcattgtgacaaagt
caaataaaccccccaaggacgacacacagtatggatcacatattgtttgacattaagcttttgccagaaaatgtt
gcatgtgtttacctcgacttgcataaatcgattagcagaaaggcatggctaataatgttgggtgggtaaaaataaat
aaataagtaaacaaaatgaagattgctgtctctctgtgcctagcctcaaagcgttcatacatcatcaccttt
aagattgctatatttgggttattttctgacaggagaaaaagatctaaagatcttttattttcatctttttggtttcttg
catgactaagaagcttaaatgttgataaaatatgactagtttgaatttacaccaagaacttctcaataaaagaa
aatcatgaatgtccacaatttcaacataccacaagagaagttaatttcttaacattgtgttctatgattattttaa
gaccttcaccaagttctgatatcttttaagacatagttcaaaattgctttgaaaaatctgtattcttgaaaatctctt
gttgtgtattagggttttaataaccagctaaaggattacctcactgagtcacagtacccctctattcagctcccaa
gatgatgtgttttgcctaccctaagagagggttttcttatttttagataattcaagtgccttagataaattatgtttctt
aagtgttatggtaaactcttttaagaaaaatttaataigtatagctgaatcttttggttaactttaaatctttatcatag
actctgtacatatgttcaaattagctgcttgcctgatgtgtgtatcatcggtgggatgacagaacaaacatatttat
gatcatgaataatgtgctttgtaaaaagattcaagttataggaagcatactctgttttttaatcatgtataatattcc
atgatacttttatagaacaattctggcttcaggaaagtctagaagcaatatttcttcaataaaaagggttttaactt
taaaaaaaaaaaaaaaaaaaaaaaaaaaaaa

FIG. 1A ii

>human SULF1 amino acid sequence--translation of ORF
MKYSCCALVLAVLGTLLGSLCSTVRSRFRGRIGQERKNIRPNIIILVLTDDQDVELG
SLQVMNKTRKIMEHGGATFINAFVTTTPMCCPSRSSMLTGKYVHNHNVYTNNENCS
SPSWQAMHEPRTFAVYLNNTGYRTAFFGKYLNEYNGSYIPPGWREWLGLIKNSRF
YNYTVCRNGIKEKHGFDYAKDYFTDLITNESINYFKMSKRMYPHRPVMMVISHAAP
HGPEDSAPQFSKLYPNASQHITPSYNYAPNMDKHWIMQYTGPMPLPIHMEFTNILQR
KRLQTLMSVDDSVRLYNMLVETGELENTYIIYTADHGYHIGQFGLVKGKSMPYDF
DIRVPFFIRGPSVEPGSIVPQIVLNIDLAPTILDIAGLDTTPDVGKSVLKLLDPEKPG
NRFRTNKKAKIWRDTFLVERGKFLRKKEESSKNIQQSNHLPKYERVKELCQQARY
QTACEQPGQKWQCIEDTSGKLRIHKCKGPSDLLTVRQSTRNLYARGFHDKDKECS
CRESGYRASRSQRKSQRQFLRNQGTTPKYKPRFVHTRQTRSLSVFEFEGEIIDINLE
EEEEELQVLQPRNIAKRHDEGHKGPRDLQASSGGNRGRMLADSSNAVGPPPTTVRV
THKCFILPNDSIHCERELYQSARAWKDHKAYIDKEIEALQDKIKNLREVRGHLKRRK
PEECSCSKQSYYNKEKGVKKQEKLKSHLHPFKEAAQEVD SKLQLFKENRRRRKKE
RKEKRRQRKGEECSLPGLTCFTHDNNHWQTAPFWNLGSFCACTSSNNNTYWCL
RTVNETHNFLFCEFATGFLEYFDMNTDPYQLTNTVHTVERGILNQLHVQLMELRSC
QGYKQCNP RPKNLDVGNKDGGSYDLHRGQLWDGWEG

FIG. 1B

FIG. 1B

>human SULF2 full length cDNA (ORF highlighted in capitals)

TGAgactcccgcatcccaaaagaagcaccagatcagcaaaaaagaagATGGGCCCCCGAGCCTCGT
GCTGTGCTTGCTGTCCGCAACTGTGTTCTCCCTGCTGGGTGGAAGCTCGGCCTTCCT
GTCGCACCACCGCCTGAAAGGCAGGTTTCAGAGGGACCGCAGGAACATCCGCCCCA
ACATCATCCTGGTGCTGACGGACGACCAGGATGTGGAGCTGGGTTCATGCAGGTG
ATGAACAAGACCCGGCGCATCATGGAGCAGGGCGGGGCGCACTTCATCAACGCCTT
CGTGACCACACCCATGTGCTGCCCCCTCACGCTCCTCCATCCTCACTGGCAAGTACGT
CCACAACCACAACACCTACACCAACAATGAGAACTGCTCCTCGCCCTCCTGGCAGGC
ACAGCACGAGAGCCGCACCTTTGCCGTGTACCTCAATAGCACTGGCTACCGGACAGC
TTTCTTCGGGAAGTATCTTAATGAATACAACGGCTCCTACGTGCCACCCGGCTGGAAG
GAGTGGGTGCGACTCCTTAAAAACTCCCGCTTTTATAACTACACGCTGTGTGCGGAACG
GGGTGAAAGAAAAGCACGGCTCCGACTACTCCAAGGATTACCTCACAGACCTCATCA
CCAATGACAGCGTGAGCTTCTTCCGCACGTCCAAGAAGATGTACCCGCACAGGCCAG
TCCTCATGGTCATCAGCCATGCAGCCCCCACGGCCCTGAGGATTCAGCCCCACAAT
ATTCACGCCTCTTCCCAAACGCATCTCAGCACATCACGCCGAGCTACAACCTACGCGC
CCAACCCGGACAAACACTGGATCATGCGCTACACGGGGCCCATGAAGCCCATCCACA
TGGAATTCACCAACATGCTCCAGCGGAAGCGCTTGACAGACCCTCATGTGCGGTGGACG
ACTCCATGGAGACGATTTACAACATGCTGGTTGAGACGGGCGAGCTGGACAACACGT
ACATCGTATACACCGCCGACCACGGTTACCACATCGGCCAGTTTGGCCTGGTGAAAG
GGAAATCCATGCCATATGAGTTTGACATCAGGGTCCCGTTCTACGTGAGGGGCCCCA
ACGTGGAAGCCGGCTGTCTGAATCCCCACATCGTCCTCAACATTGACCTGGCCCCCA
CCATCCTGGACATTGCAGGCCTGGACATACCTGCGGATATGGACGGGAAATCCATCCT
CAAGCTGCTGGACACGGAGCGGCCGGTGAATCGGTTTCACTTGAAAAAGAAGATGA
GGGTCTGGCGGACTCCTTCTTGGTGAGAGAGGCAAGCTGCTACACAAGAGAGAC
AATGACAAGGTGGACGCCAGGAGGAGAAGTTTCTGCCCAAGTACCAGCGTGTGAA
GGACCTGTGTCAGCGTGCTGAGTACCAGACGGCGTGAGCAGCTGGGACAGAAGT
GGCGATGTGTGGAGGACGCCACGGGGAAGCTGAAGCTGCATAAGTGCAAGGGCCC
CATGCGGCTGGGCGGCAGCAGAGCCCTCTCCAACCTCGTGCCCAAGTACTACGGGC
AGGGCAGCGAGGCCTGCACCTGTGACAGCGGGGACTACAAGCTCAGCCTGGCCGG
ACGCCGGAAAAAACTCTTCAAGAAGAAGTACAAGGCCAGCTATGTCCGCAGTCGCTC
CATCCGCTCAGTGGCCATCGAGGTGGACGGCAGGGTGTACCACGTAGGCCTGGGTG
ATGCCGCCCAGCCCCGAAACCTACCAAGCGGCACTGGCCAGGGGGCCCCCTGAGGA
CCAAGATGACAAGGATGGTGGGGACTTCAGTGGCACTGGAGGCCTTCCCGACTACT
CAGCCGCCAACCCCATTAAGTGACACATCGGTGCTACATCCTAGAGAACGACACAG
TCCAGTGTGACCTGGACCTGTACAAGTCCCTGCAGGCCTGGAAAGACCACAAGCTG
CACATCGACCACGAGATTGAAACCCTGCAGAACAAAATTAAGAACCTGAGGGGAAGTC
CGAGGTACCTGAAGAAAAAGCGGCCAGAAGAATGTGACTGTCACAAAATCAGCTAC
CACACCCAGCACAAAGGCCGCCTCAAGCACAGAGGCTCCAGTCTGCATCCTTTCAG
GAAGGGCCTGCAAGAGAAGGACAAGGTGTGGCTGTTGCGGGAGCAGAAGCGCAAG
AAGAACTCCGCAAGCTGCTCAAGCGCCTGCAGAACAACGACACGTGCAGCATGCC
AGGCCTCACGTGCTTCACCCACGACAACCAGCACTGGCAGACGGCGCCTTTCTGGA
CACTGGGGCCTTTCTGTGCCTGCACCAGCGCCAACAATAACACGTACTGGTGCATGA
GGACCATCAATGAGACTCACAATTTCTTCTGTGAATTTGCAACTGGCTTCCTAGA
GTACTTTGATCTCAACACAGACCCCTACCAGCTGATGAATGCAGTGAACACACTGGAC
AGGGATGTCCTCAACCAGCTACACGTACAGCTCATGGAGCTGAGGAGCTGCAAGGGT
TACAAGCAGTGTAAACCCCGGACTCGAAACATGGACCTGGGACTTAAAGATGGAGGA
AGCTATGAGCAATACAGGCAGTTTCAGCGTCGAAAGTGGCCAGAAATGAAGAGACCT
TCTTCAAATCACTGGGACAAGTGTGGGAAGGCTGGGAAGGTTAgaacaacagaggtgg

FIG. 2A i

acctcaaaaacatagaggcatcacctgactgcacaggcaatgaaaaacatgtgggtgattccagcagacctgtgctat
tggccaggaggcctgagaaagcaagcacgcactctcagtcaacatgacagattctggaggataaccagcaggagcaga
gataacttcaggaagtccattttgcccctgcttttgcttgattatacctcaccagctgcacaaaatgcatttttcgtatcaaaa
agtcaccactaaccctccccagaagctcacaaaggaaaacggagagagcgagcgagagagattccttgaaatttctc
ccaagggcgaaagtcattggaatttttaaatcataggggaaaagcagtcctgttctaaatcctctattctttgggttgacaaa
gaaggaactaagaagcaggacagaggcaacgtggagaggctgaaaacagtcagagacgttgacaatgagtcagta
gcacaaaagagatgacatttacctagcactataaaccttggtgcctctgaagaaactgccttcattgtatatgtgactattta
catgtaatcaacatgggaacttttaggggaacctataagaaatcccaattttcaggagtgggtggtcaataaacgctctgtg
gccagtgtaaaagaaaaaaaaaaaaaaaaa

FIG. 2A ii

100596 13101
T.O.T. 935207

>human SULF2 amino acid sequence--translation of ORF
MGPPSLVLCLLSATVFSLLGGSSAFLSHHRLKGRFQRDRRNIRPNILVLTDDQDVELGSM
QVMNKTRRIMEQGGAHFINAFVTPMCCPSRSSILTGKYVHNHNTYTNNENCSSPSWQ
AQHESRTFAVYLNSTGYRTAFFGKYLNEYNGSYVPPGWKEWVGLLKNSRFYNYTLCRN
GVKEKHGSDYSKDYLTDLITNDSVSFFRTSKKMYPHRPVLMVISHAAPHGPEDSAPQYS
RLFPNASQHITPSYNYAPNPDKHWIMRYTGPMKPIHMEFTNMLQRKRLQTLMSVDDSM
TIYNMLVETGELDNTYIVYTADHGYHIGQFGLVKGKSMPYEFDIRVPPFYVRGPNVEAGCL
NPHIVLNIDLAPTILDIAGLDIPADMDGKSILKLLDTERPVNRFHLKKKMRVWRDSFLVERG
KLLHKRDNDKVDAQEENFLPKYQRVKDLQCRAEYQTACEQLGQKWQCVEDATGKCLKLH
KCKGPMRLGGSRALSNLVPKYYGQGSEACTCDSGDYKLSLAGRRKKLFKKKYKASYVR
SRSIRSAIEVDGRVYHVGLGDAAQPRNLTKRHWP GAPEDQDDKDGGDFSGTGGLPDY
SAANPIKVTHRCYILENDTVQCDLDLYKSLQAWKDHKLHIDHEIETLQNKIKNLREVRGHL
KKKRPEECDCHKISYHTQHKGRCLKHRGSSLHPFRKGLQEKDQVWLLREQKRKKKLRKLL
KRLQNNDTCSMPGLTCFTHDNQHWQTAPFWTLGPFCACTSANNNTYWCMRTINETHN
FLFCEFATGFLEYFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCKGYKQCNPRTR
NMDLGLKDGGSYEQYRQFQRRKWPEMKRPSSKSLGQLWEGWEG

FIG. 2B

FIG. 2B

>mouse SULF1 full length cDNA (ORF fragments highlighted in capitals)

cttcaccttgagaaggtgaattccctaaagacatgcagtttctcaagccagaatccttgagggaaccttcaaaggactcctt
ctgcagatgttttgaaacctctgagctagaaatcgattattcaccaggatccttattcaagctcccagaactcaccgacc
aaggagcttggaagactttgcaactttgaccaagcacaATGAAGTATTCCCTCTGGGCTCTGCTGCTT
CCCCTGCTGGGCACACAGCTGCTGGGAACCCTGTGTTCCACCGTTCCGGTCCCAGAG
GTTCCGAGGAAGGATACAGCAGGAACGAAAAACATCCGACCCAACATTATTCTTGTG
CTTACCGATGATCAAGATGTGGAGCTGGGGTCCCTGCAAGTCATGAACAAAACGAGA
AAGATTATGGAACATGGGGGGGGCCACCTTCATCAATGCCTTTGTGACTACACCCATGT
GCTGCCCGTCACGGTCCCTCCATGCTCACCGGGAAGTATGTGCACAATCACAATGTCT
ACACCAACAACGAGAAGTCTCTTCCCCCTCGTGGCAGGCCATGCATGAGCCTCGG
ACTTTTGCTGTATATCTTAACAACACTGGCTACAGAACAGCCTTTTTTGGAAAATACCT
CAATGAATATAATGGCAGCTACATCCCCCTGGGTGGCGAGAATGGCTTGGATTAATC
AAGAATTCTCGTTTCTATAATTACACTGTTTGTGCAATGGCATCAAAGAAAAGCATGG
ATTTGATTATGCAAAGGACTACTTCACAGACTTAATCACTAACGAGAGCATTAACTACTT
CAAAATGTCTAAGAGAATGTATCCCCATAGGCCCGTTATGATGGTGATCAGCCACGCT
GCGCCCCACGGCCCCGAGGACTCAGCCCCACAGTTTTTCTAAACTGTACCCCAATGCT
TCCCAACACATAACTCCTAGTTATAACTATGCACCAAATATGGATAAACACTGGATTATG
CAGTACACAGGACCAATGCTGCCCATCCACATGGAATTTACAAACATTCTACAGCGCA
AAAGGCTCCAGACTTTTGATGTCAGTGGATGATTCTGTGGAGAGGCTGTATAACATGCT
CGTGGAGACGGGGGAGCTGGAGAATACTTACATCATTTACACCGCCGACCATGGTTA
CCATATTGGGCAGTTTGGACTGGTCAAGGGGAAATCCATGCCATATGACTTTGATATTC
GTGTGCCTTTTTTTATTCTGTTGTTCCAAGTGTAGAACCAGGATCAATAGTCCCACAGATC
GTTCTCAACATTGACTTGGCCCCACGATCCTGGATATTGCTGGGCTCGACACACCTC
CTGATGTGGACGGCAAGTCTGTCTCAAACCTTCTGGACCCAGAAAAGCCAGGTAACA
GGTTTTCGAACAAACAAGAAGGCCAAAATTTGGCGTGATACATTCCTAGTGGAAGAGG
CAAATTTCTACGTAAAGAAGGAAGAATCCAGCAAGAATATCCAACAGTCAAATCACTTGC
CCAATATGAACGGGTCAAAGAAGTATGCCAGCAGGCCAGGTACCAGACAGCCTGTG
AACAAACCGGGGCAGAAGTGGCAATGCATTGAGGATACATCTGGCAAGCTTCGAATTC
ACAAGTGTAAGGACCCAGTGACCTGCTCACAGTCCGGCAGAGCACGCGGAACCTC
TACGCTCGCGGCTTCCATGACAAAGACAAAGAGTGCAGTTGTAGGGAGTCTGGTTAC
CGTGCCAGCAGAAGCCAAAGAAAGAGTCAACGGCAATTCTTGAGAAACCAGGGGAC
TCCAAAGTACAAGCCCAGATTTGTCCATACTCGGCAGACACGTTCTTGTCCGTCGAA
TTTGAAGGTGAAATATATGACATAAATCTGGAAGAAGAAGAAGATTGCAAGTGTTGCA
ACCAAGAAACATTGCTAAGCGTCATGATGAAGGCCACAAGGGGGCCAAGAGATCTCCA
GGCTTCCAGTGGTGGCAACAGGGGCAGGATGCTGGCAGATAGCAGCAACGCCGTGG
GCCACCTACCACTGTCCGAGTGACACACAAGTGTTTTATTCTTCCCAATGACTCTATC
CATTGTGAGAGAGAACTGTACCAATCGGCCAGAGCGTGGAAGGACCATAAGGCCTAC
ATTGATAAAGAGATTGAAGTTCTACAAGATAAAATTAAGAATTTAAGGGAAGTGAGGGG
ACACCTAAAGAAAAGGAAACCTGAGGAGTGTAGCTGTGGTGACCAGAGCTATTACAA
CAAAGAGAAAAGGTGTCAAACGACAGGAGAAGCTAAAGAGTCACCTTCACCCCTTCAA
GGAGGCTGCTGCCAGGAGGTGGATAGCAAACCTTCAGCTCTTCAAGGAGCATCGGA
GGAGGAAGAAGGAGAGGAAGGAGAAGAAACGGCAGAGGAAGGGAGAGGAGTGTAG
CCTGCCTGGCCTTACCTGCTTCACCCATGACAACAACCACTGGCAGACTGCCCCATT
CTGGAACCTTGGGATCTTTCTGTGCCTGCACAAGTTCTAACAACAATACCTACTGGGTG
TTGCGTACAGTCAACGAGACGCACAATTTCTGTTTTGTGAGTTTGCTACTGGCTTTC
TGGAATATTTGACATGAATACGGATCCTTATCAGCTCACAAATACAGTACACACAGTA

FIG. 3A i

GAACGAGGCATTTTGAATCAGCTACACGTACAACCTAATGGAGCTCAGAAGCTGTCA
AGGATATAAGCAGTGCAACCCAAGACCTAAGAATCTTGATGTTGGAAATAAAGATG
GAGGAAACTATGACCCGCACAGAGGACAGTTATGGGATGGATGGGAAGGTTAGTC
TTTCCAATGTTACTTCAGACACCAGCTGGCAAGGCCTGGAGGAGTTATCCGGTGC
AAGCGACATCGATGAGTACAGGTCTAACCCTAGACTAAGTCTGGAGGACTGGACT
AACTACCTGAGGGCTGTCTACAGAGCCTTTGCACTGCTGAACAGTCACCCTGATC
CAAACAAAGCAAATGGGACTCCAACCACACAAGGTGGTGACTTCCTGGTCACCTC
TGCTGAGCGCTTGGTGCCAGCAGAGATGGCTTCTGCAGAATCAGGTGAAGACCC
AAGTCATGTGGTTGGGGAAACACCTCCTTTGACCTTGCCAGTCAACCTCCAAACC
CTGCATCTGAACAGACCAACGTTAAGTCCAGAGAGAAAACCTTGAATGGGATAATGA
CATTCCAGAAGTGAATCATTTGAATTCTGAACACTGGAGAAAAACTGAGAAGCAGA
TAGGATGGGAGGAGCTGCATCATCCTGAAGGTGACGTCGTCAGTGGCAATGGTAT
GACAGAGCTGCTGCCCCAGTCTCATCTTGGGCATCAGCTCACCAGTCAGCACCA
ACAAAAATGTTCCCAGGATGTGGAGACAGAGAAGGATGCTTTTGAAGATCAATTG
CGTCCTCTTGTCCACTCTGACAGAACTCCGGTTCATC

FIG. 3A ii

100596 101

>mouse SULF1 amino acid sequence--translation of ORF
MKYSLWALLLP LLGTQLLGTL CSTVRSQRFRGRIQQERKNIRPN IILVLTDDQDVELGS
LQVMNKTRKIMEHGGATFINAFVTT PMCCPSRSSMLTGKYVHNHNVYTNNENCSSP
SWQAMHEPRTFAVYLNNTGYRTAFFGKYLNEYNGSYIPPGWREWLG LIKNSRFYNY
TVCRNGIKEKHGFDYAKDYFTDLITNESINYFKMSKRMYPHRPVMMVISHAAPHGPE
DSAPQFSKLYPNASQHITPSYNYAPNMDKHWIMQYTG PMLPIHMEFTNILQRKRLQT
LMSVDDSVRLYNMLVETGELENTYIIYTADHGYHIGQFGLVKGKSMPYDFDIRVPFFI
RGPSVEPGSIVPQIVLNIDLAPTILDIAGLDTPPDVDGKSVLKLLDPEKPGNRFRTNKK
AKIWRDTFLVERGKFLRKKEESSKNIQQSNHLPKYERVKELCQQARYQTACEQPGQ
KWQCIEDTSGKLRIHKCKGPSDLLTVRQSTRNLYARGFHDKDKECSCRESGYRASR
SQRKSQRQFLRNQGT PKYKPRFVHTRQTRSLSEFEGEIYDINLEEEEEELQVLQPRN
IAKRHDEGHKGPRDLQASSGGNRGRMLADSSNAVGPPTTVRVTHKCFILPND SIHCE
RELYQSARAWKDHKAYIDKEIEVLQDKIKNLREVRGHLKKRKPEECSCGDQSYYNKE
KGVKRQEKLKSHLHPFKEAAAQEVDSKLQLFK EHRRRKKERKEKKRQRKGEECSLP
GLTCFTHDNNHWQTAPFWNLGSFCACTSSNNNTYWVLRTVNETHNFLFCEFATGFL
EYFDMNTDPYQLTNTVHTVERGILNQLHVQLMELRSCQGYKQCNP RPKNLDVGNKD
GGNYDPHRGQLWDGWEG

FIG. 3B

1005966-12401

>mouse SULF-2 cDNA (ORF in capital letters)
ggacgcgtgggacgcgcgtggggtctgggcaacgcttctgcttgccttgagctcaactaatttctcagagagcttcgg
agacgcgtgggaaggtcccaggcgcgtgggcagttctctccgcgcatctagctggggatcgccccgagccggcgtctc
caatgatcctgaggaagaggggaaggaatcccatcctcacgacaccacctcgccctctgcatccaggaagaagca
aaggaccagcaagccacgccaATGGCACCCCCCTGGCCTGCCACTATGGCTGCTGTCCAC
CGCTCTCCTCTCCCTGCTGGCTGGCAGCTCGGCCTTCTCTCCCATCCCCGCCT
GAAGGGACGCTTCCAGAGGGACCGCAGGAACATCCGGCCCCAACATCATCTTGGT
GCTTACGGATGACCAGGATGTGGAGCTGGGCTCCATGCAAGTGATGAACAAGACA
AGGCGTATCATGGAGCAGGGCGGGGCGCACTTCATCAATGCCTTCGTGACTACAC
CAATGTGCTGTCCGTCTCGCTCCTCCATTCTCACC GGCAAGTACGTCCACAACCA
CAACACCTACACCAACAATGAGAATTGTTCTCGCCCTCCTGGCAGGCCCCAGCAC
GAGAGCCGCACCTTCGCCGTGTATCTCAACAGCACAGGCTACCGGACAGCTTTCT
TCGGAAAATACCTCAATGAGTACAACGGCTCATACGTGCCGCCCCGGCTGGAAGGA
GTGGGTGGGCCTACTTAAGAACTCCCGCTTTTATAACTACACACTCTGCCGGAATG
GGGTGAAGGAGAAACATGGCTCAGACTACTCCACGGATTACCTCACGGATCTCAT
CACCAATGACAGTGTGAGCTTCTTCCGAACATCCAAGAAGATGTACCCACACAGG
CCCGTGCTCATGGTCATCAGCCACGCGGCTCCCCACGGCCCCGAGGACTCGGC
ACCGCAGTACTCACGGCTCTTCCCCAATGCGTCCCAGCACATCACACCGAGTTAC
AACTATGCACCCCAACCCAGACAAGCATTGGATCATGCGCTACACGGGACCCATGA
AGCCCATTCACATGGAATTCACCAACATGCTACAACGCAAACGCCTACAGACCCTC
ATGTCTGTGGATGACTCCATGGAGACGATCTATGACATGCTGGTGGAGACGGGGG
AGCTGGACAACACGTACATCCTGTACACCGCCGACCACGGCTACCACATTGGCCA
GTTTGGGCTGGTGAAGGGCAAGTCTATGCCGTATGAATTCGACATCAGAGTCCCG
TTCTACGTGAGGGGCCCCAACGTGGAAGCTGGCTCTCTGAACCCCCACATTGTC
CTCAACATTGACCTGGCCCCCACCATACTGGATATCGCTGGACTGGACATCCCTG
CAGACATGGACGGGAAGTCTATTCTCAAACACTACTGGACTCAGAGCGGCCAGTGAA
CCGGTTCCACTTGAAAAAGAAGCTGAGGGTCTGGCGAGACTCCTTCCTGGTGGA
GAGAGGCCAACTGCTCCACAAGAGGGAGGGTGACAAAGTGAATGCCCAGGAGGA
GAACTTCCTGCCCAAGTACCAGCGCGTGAAGGACCTGTGTGTCAGCGAGCTGAGTA
CCAGACAGCATGCGAACAGCTGGGGCAGAAGTGGCAGTGTGTGGAGGACGCTT
CTGGGACGCTGAAGCTGCACAAATGTAAAGGCCCCATGCGGTTTGGTGGCGGCG
GTGGCAGCAGAGCCCTCTCCAACCTGGTGCCCAAGTATGACGGCCAGAGCAGCG
AGGCCTGCAGCTGTGACAGTGGCGGTGGAGGGGACTACAACTGGGCCTGGCT
GGACGCCGTAAAGCTCTTTAAGAAAAAGTATAAGACCAGCTATGCCCGGAACCGCT
CCATCCGTTCCGTGGCCATCGAGGTGGACGGTGAGATATACCACGTAGGCTTGGA
TACTGTGCCTCAGCCCCGCAACCTTAGCAAGCCGCACTGGCCAGGGGGCCCCCTGA
AGACCAAGATGACAAGGATGGTGGCAGTTTCAGTGGTACTGGTGGCCTTCCAGAT
TATTCTGCCCCCAATCCCATCAAAGTGACCCATCGGTGCTACATCCTTGAGAATGA
CACAGTCCAGTGCGACTTGGACCTGTACAAGTCCCTGCAGGCTTGGAAGACCA
CAAGCTGCACATCGACCATGAGATCGAAACCCTGCAGAACAAAATTAAGAACCTTC
GAGAAGTCAGGGGTACCTGAAGAAGAAGCGACCGGAAGAATGTGACTGCCATA
AAATCAGTTACCACAGCCAACACAAAGGCCGTCTCAAGCACAAAGGCTCCAGCCT
GCACCCTTTCAGGAAGGGTCTGCAGGAGAAGGACAAGGTGTGGCTGCTGCGGG
AGCAGAAACGCAAGAAGAACTGCGCAAGCTGCTCAAACGGCTGCAGAACAAACG
ATACGTGCAGCATGCCCGGCCTCACGTGCTTTACCCACGACAACCACCACTGGCA
GACGGCGCCACTCTGGACGCTGGGGCCGTTCTGCGCCTGCACCAGCGCCAACA
ACAACACGTACTGGTGCTTGAGGACCATAAATGAGACCCACAACCTTCCTCTTCTGC
GAATTTGCAACCGGCTTCATAGAATACTTTGACCTCAGTACAGACCCCTACCAGCT
GATGAACGCGGTGAACACACTGGACAGGGACGTCTTAACCAACTGCACGTGCA
GCTCATGGAGCTAAGGAGCTGTAAAGGCTACAAGCAGTGCAACCCCCGGACCCG
CAACATGGACCTGGGGCTTAGAGACGGAGGAAGCTATGAACAATACAGGCAGTTT
CAGCGTCGAAAATGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGGGACAGC
TATGGGAAGGTTGGGAAGGCTAAgcgccatagagagaggaacctccaaaaccaggggcctcgtgtg

FIG. 4Ai

gctgcccaggccatgcaaaaaacacccgattcccagaagatgaatgttggaactgggagacctgacagaaggcagg
gctgctctgggacaggaaatcctggaggacagcgctggactttccgatgctcagtttcttgccctgcttgctctggatca
aacctcactggctgctctgggatgctgctcacacctggagtgctcacccttcagagggtcacaaagacaaagga
actaatccatggacacttcctccagagatggaaattgctgggattcgcccactcctcccctgcacccctccccagtc
ctagggaagcaagctgttttaaccttcttactcttggagaaagcacggacatcccagggtgctgtcaacctcacagtctga
caaagtctatagcacaaaacagtaccattcaccagggtggtgacctggctggctcagaagctgccttcaccacatacat
gaccgctcacacgtaaccaacacagggaattgtaggggaatctcactaatatgaaatcccgcittcaagagtcgcggtg
tcaataaacgctgtggctaggatcaaggataatccctgagcttcagacatttctcctgccgggattcggttccttgtatcc
atcccagaactgatgttttctaaggtaaccgaaacccaagtgatgtgtctctgtgttttaatgacattgtattgtaaagttt
gtagtataagtaccatcttacagtgttctgccccagccaatgtctagctattggtatgaaaaaaaaaattcttgaattttg
taaaaaaaaaaaaaa

FIG. 4A ii

FIG. 4A ii

>mouse SULF2 amino acid sequence--translation of ORF
MAPPG LPLWLLSTALLSLLAGSSAFLSHPR LKGRFQRDRRNIRPN IILVLTDDQDVELGSM
QVMNKTRRIMEQGGAHFINAFVTTPMC CPSRSSILT GKYVHNHNTYTN NENCSSPSWQ
AQHESRTFAVYLNSTGYRTAFFGKYLNEYNGSYVPPGWKEWVG LLLKNSRFYNYTL CRN
GVKEKHGSDYSTDYLTDLITNDSVSFFRTSKKMYPHRPVLMVISHAAPHG PEDSAPQYS
RLFPNASQHITPSYNYAPNPDKHWIMRYTGPMKPIHMEFTNMLQRKRLQTLMSVDDSM
ETIYDMLVETGELDNTYILYTADHG YHIGQFGLVKGKSMPYEFDIRVPFYVRGPNVEAGSL
NPHIVLNIDLAPTILDIAGLDIPADMDGKSILKLLDSE RPNRFLKLLRVWRDSFLVERG
KLLHKREGDKVNAQEENFLPKYQRVKDLCQRAEYQTACEQLGQKWQCVEDASGTLKL
HKCKGPMRFGGGGGSRALSNLVPKYD GQSSEACSCDSGGGGDYKLGLAGRRKLFKKK
YKTSYARNRSIRSVAIEVDGEIYHVGLDTVPQPRNL SKPHWPGAPEDQDDKDGGSFSGT
GGLPDYSAPNPIKVTHRCYILENDTVQCDLDLYKSLQAWKDHKLHIDHEIETLQNKIKNLR
EVRGHLKKKRPEECDCHKISYHSQHKGR LKHGSSLHPFRKGLQE KDKVWLLREQKRK
KKLRKLLKRLQNNDTC SMPGLTCFTHDNHHWQTAPLWTLGPFC ACTSANNNTYWCLRT
INETHNFLFC EFATGFIEYFDLSTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCKGYKQC
NPRTRNMDLGLRDGGSYEQYRQFQRRKWPEMKRPSSKSLGQLWEGWEG-

FIG. 4B

FIG. 4B

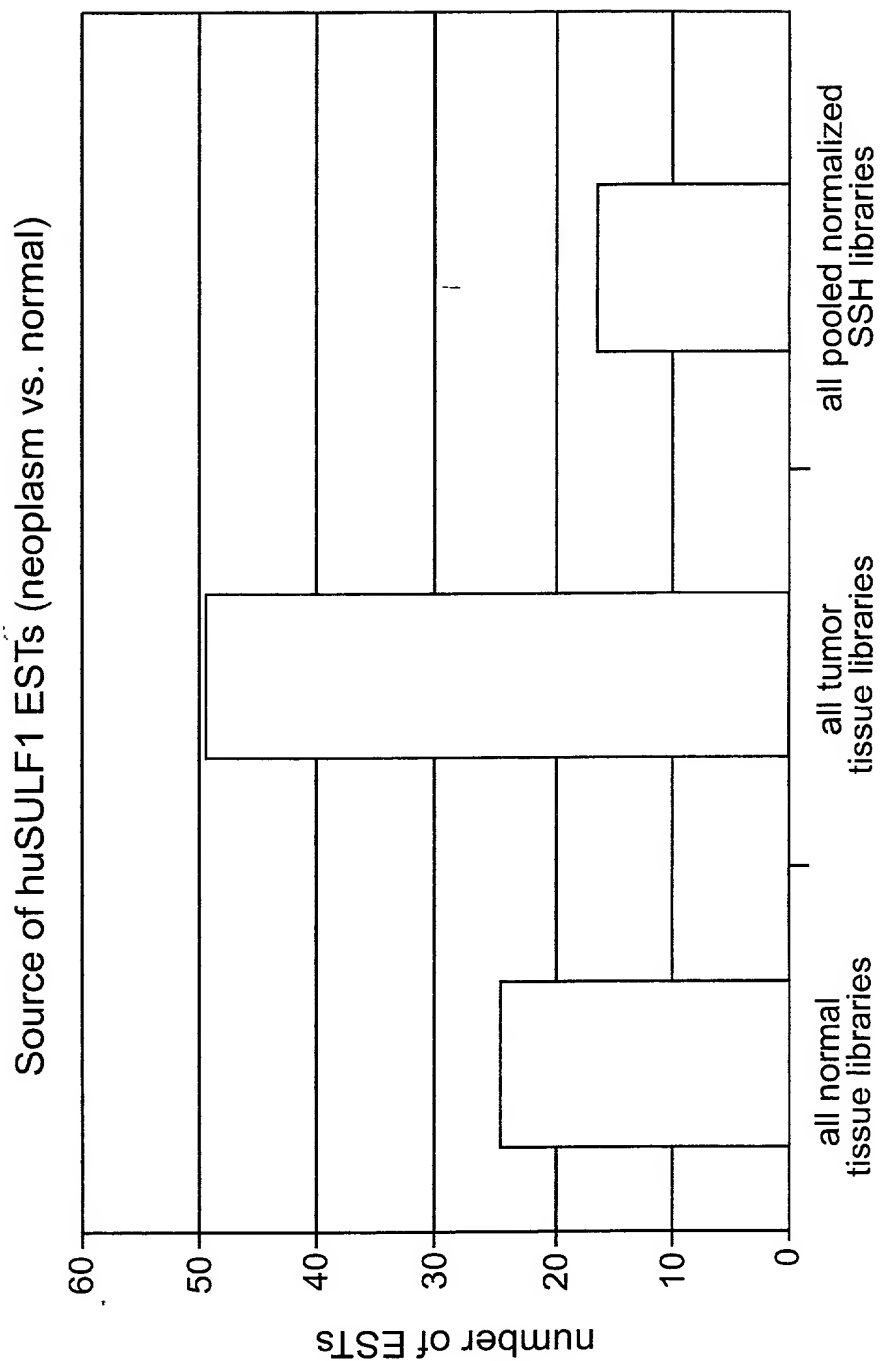


FIG. 5

Source of huSULF1 ESTs (detailed)

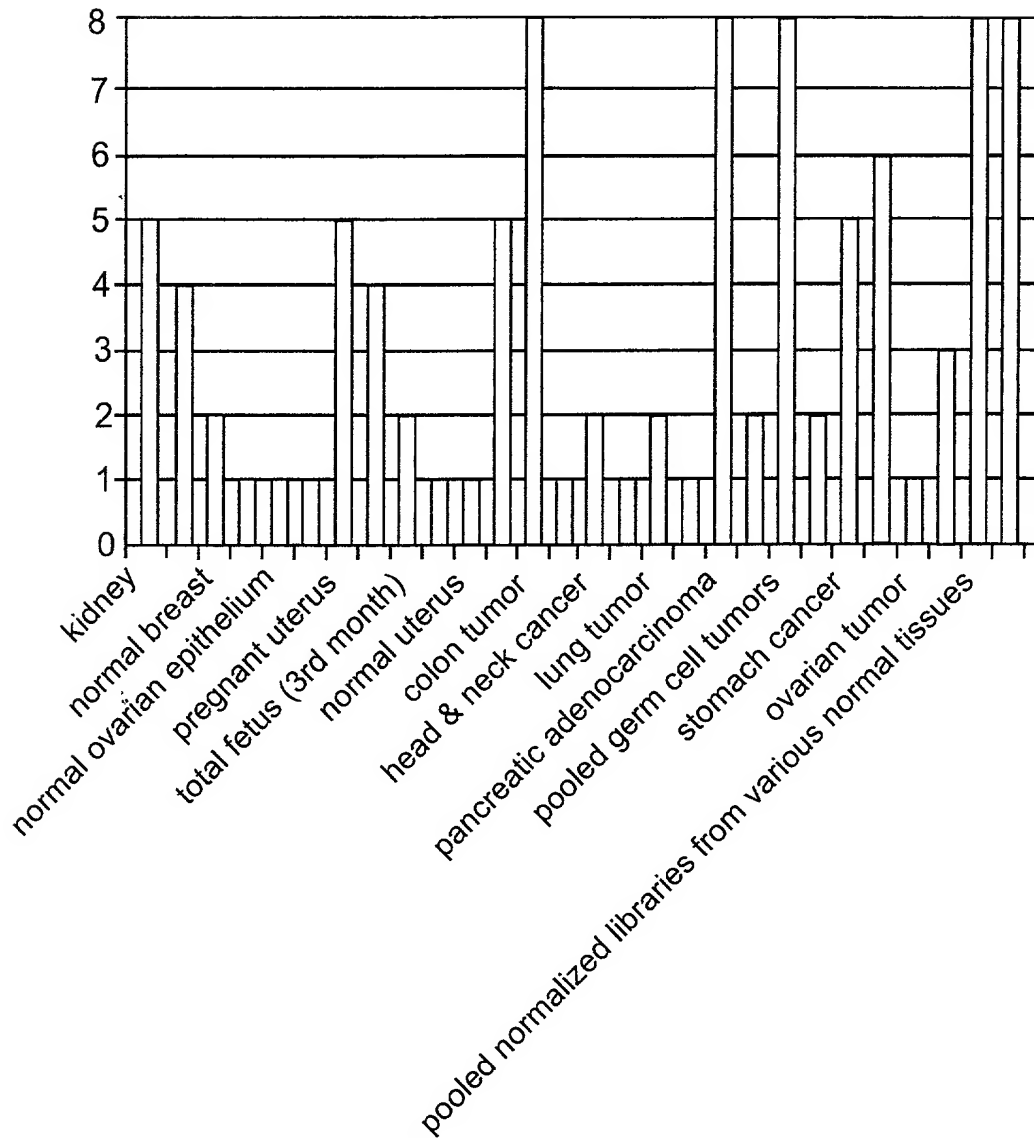


FIG. 6

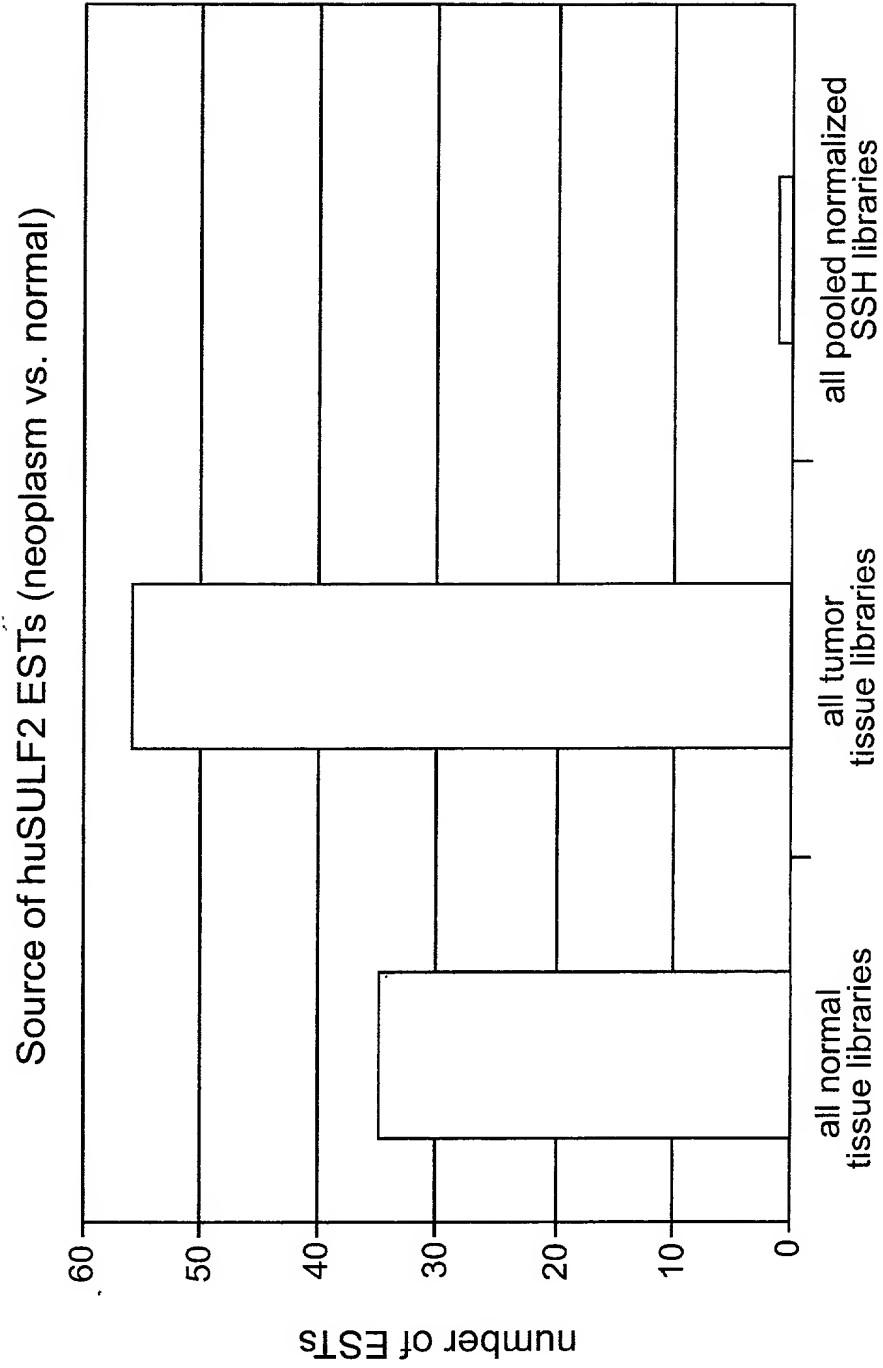


FIG. 7

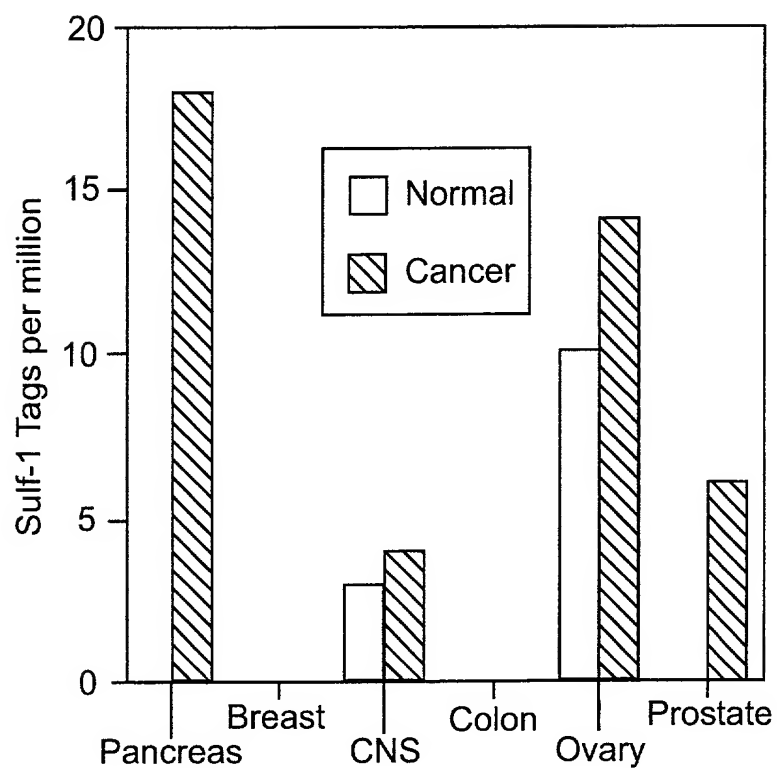


FIG. 8

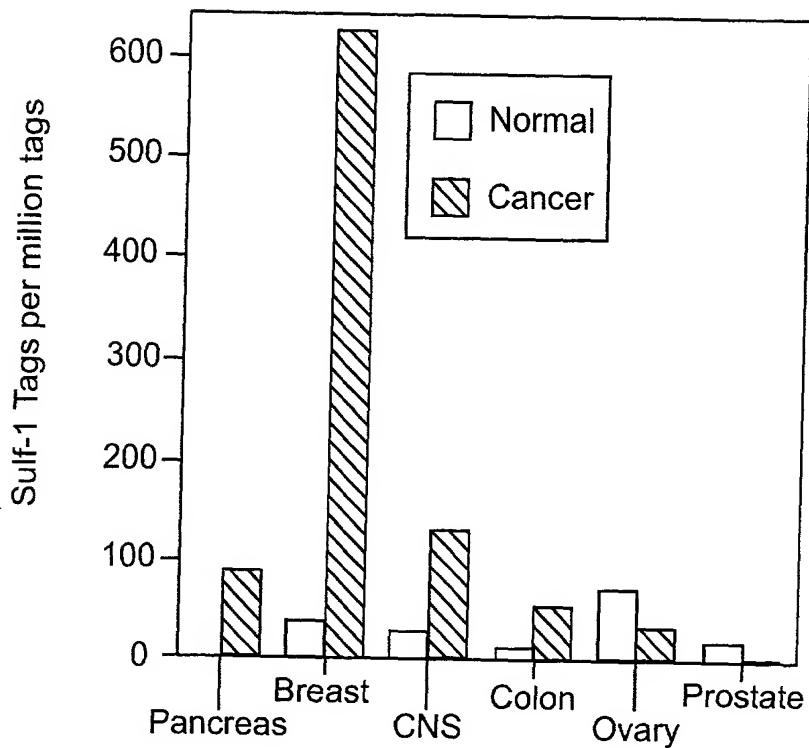


FIG. 9

> human SULF2 full length cDNA (ORF is highlighted in capitals and a 5' inframe stopcodon is underscored)

ggcacgagggccatttctggacaacagctgctattttcaacttgagcccaagttaatttctcggggagtttctcgggcgcgcacagagctcggttggcctgcgattgagctgcgggtcgcgccgagcgccgctctccaatggcaaagtgtgtgtggctggaggcgagcgcgaggctttcggaagggcagtcgagtggttcgagaccggggcgagtcctgtgaaagcagataaaagaaaaatttattaacgtgtcattacgagggggagcgcccgccggggctgtcgactccccgcggaacatttggctccctccagctcctagagaggagaagaagaaagcggaagagggcagattcacgtcggttccagccaagtggacctgatcgatggccctcctgaatttatcacgatatttgatttattagcgatgccccctggttgtgtgttacgcacacacacgtgcacacaaggctctggctcgttccctccctcggttccagctcctggcgaaatccacatctgtttcaactctccgccgagggcgagcaggagcgagagtggtcgaatctgcgagtgaaagaggagcgagggaaaagaacaaagccacagacgcaacttgagactcccgcatcccaaagaagcaccagatcagcaaaaaaagaagATGGGCCCCCGAGCCTCGTGCTGTGCTTGCTGTCCGCAACTGTGTTCTCCCTGCTGGGTGGAAGCTCGGCCTTCCTGTGCGACCAACGCCTGAAAGGCAGGTTTTCAGAGGGACCGCAGGAACATCCGCCCCAACATCATCCTGTGCTGACGGACGACCAGGATGTGGAGCTGGGTTCATGCAGGTGATGAACAAGACCCGGCGCATCATGGAGCAGGGCGGGGCGCACTTCATCAACGCCTTCGTGACCACACCCATGTGCTGCCCCCTCACGCTCCTCCATCCTCACCGGCAAGTACGTCCACAACCAACACCTACACCAACAATGAGAACTGCTCCTCGCCCTCCTGGCAGGCACAGCACGAGAGCCGACCTTTGCCGTGTACCTCAATAGCACTGGCTACCGGACAGCTTTCTTCGGGAAGTATCTTAATGAATACAACGGCTCCTACGTGCCACCCGGCTGGAAGGAGTGGGTGCGACTCCTTAAAACTCCCGCTTTTATAACTACACGCTGTGTGCGAACGGGGTGAAAGAGAAGCACGGCTCCGACTACTCCAAGGATTACCTCACAGACCTCATCACC AATGACAGCGTGAGCTTCTTCCGCACGTCCAAGAAGATGTACCCGCACAGGCCAGTCTCATGGTCATCAGCCATGCAGCCCCCACGGCCCTGAGGATTCAGCCCCACAATATTCACGCCTCTTCCCAAACGCATCTCAGCACATCACGCCGAGCTACAACCTACGCGCCAAACCCGGACAAACACTGGATCATGCGCTACACGGGGCCCATGAAGCCCATCCACATGGAATTCACCAACATGCTCCAGCGGAAGCGCTTGACAGACCCTCATGTCCGTGGA CGACTCCATGGAGACGATTTACAACATGCTGGTTGAGACGGGGCGAGCTGGACAACA CGTACATCGTATACACCGCCGACCACGGTTACCACATCGGCCAGTTTGGCCTGGTGAAAGGGAAATCCATGCCATATGAGTTTGACATCAGGGTCCCCTTCTACGTGAGGGGC CCAAACGTGGAAGCCGGCTGTCTGAATCCCCACATCGTCTCAACATTGACCTGGC CCCCACCATCCTGGACATTGCAGGCCTGGACATACCTGCGGATATGGACGGGAAAT CCATCCTCAAGCTGCTGGACACGGAGCGGCCGGTGAATCGGTTTCACTTGAAAAAG AAGATGAGGGTCTGGCGGGACTCCTTCTTGGTGGAGAGAGGCAAGCTGCTACACA AGAGAGACAATGACAAGGTGGACGCCAGGAGGAGAACTTTCTGCCCAAGTACCA GCGTGTGAAGGACCTGTGTCAGCGTGCTGAGTACCAGACGGCGTGAGCAGCTG GGACAGAAGTGGCAGTGTGTGGAGGACGCCACGGGGAAGCTGAAGCTGCATAAGT GCAAGGGCCCCATGCGGCTGGGCGGCAGCAGAGCCCTCTCCAACCTCGTGCCCA AGTACTACGGGCAGGGCAGCGAGGCCTGCACCTGTGACAGCGGGGACTACAAGCT CAGCCTGGCCGGACGCCGGAaaaaactCTTCAAGAAGAAGTACAAGGCCAGCTATG TCCGCAGTCGCTCCATCCGCTCAGTGGCCATCGAGGTGGACGGCAGGGTGTACCA CGTAGGCCTGGGTGATGCCGCCAGCCCCGAAACCTCACCAAGCGGCACTGGCCA GGGGCCCTGAGGACCAAGATGACAAGGATGGTGGGGACTTCAGTGGCACTGGAG GCCTTCCCGACTACTCAGCCGCCAACCCTTAAAGTGACACATCGGTGCTACATCC TAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTACAAGTCCCTGCAGGCCTGG AAAGACCACAAGCTGCACATCGACCACGAGATTGAAACCCTGCAGAACAAAATTAAG AACCTGAGGGAAGTCCGAGGTACCTGAAGAAAAAGCGGCCAGAAGAATGTGACT GTCACAAAATCAGCTACCACACCCAGCACAAAGGCCCGCTCAAGCACAGAGGCTCC AGTCTGCATCCTTTTCAGGAAGGGCCTGCAAGAGAAGGACAAGGTGTGGCTGTTGC GGGAGCAGAAGCGCAAGAAGAACTCCGCAAGCTGCTCAAGCGCCTGCAGAACAA

FIG. 10A i

1005966-1-101

CGACACGTGCAGCATGCCAGGCCTCACGTGCTTCACCCACGACAACCAGCACTGG
CAGACGGCGCCTTTCTGGACACTGGGGCCTTTCTGTGCCTGCACCAGCGCCAACA
ATAACACGTACTGGTGCATGAGGACCATCAATGAGACTCACAATTTCTCTTCTGTG
AATTTGCAACTGGCTTCCTAGAGTACTTTGATCTCAACACAGACCCCTACCAGCTGA
TGAATGCAGTGAACACACTGGACAGGGATGTCCTCAACCAGCTACACGTACAGCTC
ATGGAGCTGAGGAGCTGCAAGGGTTACAAGCAGTGTAACCCCCGGACTCGAAACA
TGGACCTGGGACTTAAAGATGGAGGAAGCTATGAGCAATACAGGCAGTTTCAGCGT
CGAAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGGGACAACGTGTTGGG
AAGGCTGGGAAGGTTAAgaaacaacagaggtggacctccaaaaacatagaggcatcacctgactgcacag
gcaatgaaaaaccatgtgggtgatttccagcagacctgtgctattggccaggaggcctgagaaagcaagcagcactct
cagtcaacatgacagattctggaggataaccagcaggagcagagataacttcaggaagtcattttgcccctgctttgct
ttggattatacctcaccagctgcacaaaaatgcatttttcgtatcaaaaaagtcaccactaacctccccagaagctcaca
aggaaaaacggagagagcgagcgagagagatttccttgaaatttctccaagggcgaaagtcattggaattttaaatca
taggggaaaagcagtcctgttctaaatccttattctttgggttgcacaaagaaggaaactaagaagcaggacagagggc
aacgtggagaggctgaaaacagtcagagacgtttgacaatgagtcagtagcacaaaagagatgacatttacctagca
tataaacctggttgctctgaagaaactgccttcattgtatatagtgactattacatgtaataacatgggaacttttagggg
aacctaataagaaatccaatttcaggagtggtggtgtcaataaacgctctgtggccagtgtaaaagaaaaaaaaaaa
aaattgtggacatttctgttctgtccagataccatttctcctagatttcttggatgtcccagaactgatgttttttaaggtagt
gaaaagaaatgaagtgtatgtcccaagtttgatgaaactgtattgtaaaaaaattttgagttaagtattgtcataca
gtgttcaaaacccagccaatgaccagcagttggtatgaagaaccttgacattttgtaaaggccatttctggggaaaaa
aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa

FIG. 10A ii

>human SULF2 protein (translation of ORF)

MGPPSLVLCLLSATVFSLLGGSSAFLSHHRLKGRFQRDRRNIRPNILVLTDDQDVELGS
MQVMNKTRRIMEQGGAHFINAFVTTMCCPSRSSILTGYVHNHNTYTNNENCSSPS
WQAQHEsrTFAVYLNSTGYRTAFFGKYLNEYNGSYVPPGWKEWVGLLKNSRFYNYTL
CRNGVKEKHGSDYSKDYLTDLITNDSVSFFRTSKKMYPHRPVLMVISHAAPHGPEDSA
PQYSRLFPNASQHITPSYNYAPNPDKHWIMRYTGPMKPIHMEFTNMLQRKRLQTLMSV
DDSMETIYNMLVETGELDNTYIVYTADHGYHIGQFGLVKGKSMPYEFDIRVPFYVRGPN
VEAGCLNPHIVLNIDLAPTILDIAGLDIPADM DGKSILKLLDTERPVNRFHLKKKMRVWRD
SFLVERGKLLHKRDNDKVDAQEENFLPKYQRVKDLCQRAEYQTACEQLGQKWQCVED
ATGKLKLHKCKGPMRLGGSRALSNLVPKYQGQSEACTCDSGDYKLSLAGRRKKLFKK
KYKASYVRSRSIRSVAIEVDGRVYHVGLGDAAQPRNLTKRHWPGAPEDQDDKDGGDF
SGTGGLPDYSAANPIKVTHRCYILENDTVQCCLDLYKSLQAWKDHKLHIDHEIETLQNKI
KNLREVRGHLKKKRPEECDCHKISYHTQHKGRCLKHRGSSLHPFRKGLQEKDQVWLLR
EQKRKKKLRKLLKRLQNNDTCSMPGLTCFTHDNQHWQTAPFWTLGPFCACTSANNNT
YWCMRTINETHNFLFCFATGFLEYFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELR
SCKGYKQCNPRTRNMDLGLKDGGSYEQYRQFQRRKWPEMKRPSSKSLGQLWEGWE
G

FIG. 10B

1006966-100101

>mouse SULF2 full length cDNA (ORF highlighted in capitals)

ggcggcggagatcctgagggagaggggaaggaatccccatcctcagcaccacctcggcctctgcatccaggaagaa
gcaaaggaccagcaagccacgccaATGGCACCCCCTGGCCTGCCACTATGGCTGCTGTCCAC
CGCTCTCCTCTCCCTGCTGGCTGGCAGCTCGGCCTTCCTCTCCCATCCCCGCCTGA
AGGGACGCTTCCAGAGGGACCGCAGGAACATCCGGCCCCAACATCATCTTGGTGCTT
ACGGATGACCAGGATGTGGAGCTGGGCTCCATGCAAGTGATGAACAAGACAAGGCG
TATCATGGAGCAGGGCGGGGCGCACTTCATCAATGCCTTCGTGACTACACCAATGTG
CTGTCCGTCTCGCTCCTCCATTCTCACCGGCAAGTACGTCCACAACCACAACACCTA
CACCAACAATGAGAATTGTTCTCGCCCTCCTGGCAGGCCCCAGCACGAGAGCCGCA
CCTTCGCCGTGTATCTCAACAGCACAGGCTACCGGACAGCTTTCTTCGGAAAATACCT
CAATGAGTACAACGGCTCATACGTGCCGCCCGGCTGGAAGGAGTGGGTGGCCTAC
TTAAGAACTCCCGCTTTTATAACTACACACTCTGCCGGAATGGGGTGAAGGAGAAACA
TGGCTCGGACTACTCCACGGATTACCTCACGGATCTCATCACCATGACAGTGTGAG
CTTCTTCCGAACATCCAAGAAGATGTACCCACACAGGCCCGTGCTCATGGTCATCAG
CCACGCGGCTCCCCATGGCCCCGAGGACTCAGCACCCCAGTACTCACGGCTCTTCC
CCAATGCGTCCCAGCACATCACACCGAGTTAGAACTATGCACCCAACCCAGACAAGC
ATTGGATCATGCGCTACACGGGACCCATGAAGCCCATTCACATGGAATTCACCAACAT
GCTACAACGAAAACGCCTACAGACCCTCATGTCTGTGGATGACTCCATGGAGACGAT
CTATGACATGCTGGTGGAGACGGGGGAGCTGGACAACACGTACATCCTGTACACCGC
CGACCACGGCTACCACATTGGCCAGTTTGGGCTGGTGAAGGGCAAGTCTATGCCGTA
TGAATTGACATCAGAGTCCCGTTCTACGTGAGGGGGCCCCAACGTGGAAGCTGGCT
CTCTGAACCCCCACATTGTCTCAACATTGACCTGGCCCCCACCATACTGGATATCGC
TGGACTGGACATCCCTGCAGACATGGACGGGAAGTCTATTCTCAAACACTACTGGACTC
AGAGCGGCCAGTGAACCGGTTCCACTTGAAAAAGAAGCTGAGGGTCTGGCGAGACT
CCTTCCTGGTGGAGAGAGGCAAACCTGCTCCACAAGAGGGAGGGTGACAAAGTGAAT
GCCCAGGAGGAGAACTTCCTGCCCAAGTACCAGCGCGTGAAGGACCTGTGTGAGCG
AGCTGAGTACCAGACAGCATGCGAACAGCTGGGGCAGAAAGTGGCAGTGTGTGGAGG
ACGCTTCTGGGACGCTGAAGCTGCACAAATGTAAAGGCCCCATGCGGTTTGGTGGC
GGCGGTGGCAGCAGAGCCCTCTCCAACCTGGTGCCCAAGTATGACGGCCAGAGCA
GCGAGGCCTGCAGCTGTGACAGTGGCGGTGGAGGGGACTACAAACTGGGCCTGGC
TGGACGCCGTAAGCTCTTTAAGAAAAAGTATAAGACCAGCTATGCCCGGAACCGCTC
CATCCGTTCCGTGGCCATCGAGGTGGACGGTGAGATATACCACGTAGGCTTGGATAC
TGTACCTCAGCCCCGCAACCTTAGCAAGCCGCACTGGCCAGGGGGCCCTGAAGACC
AAGATGACAAGGATGGTGGCAGTTTCAGTGGTACTGGTGGCCTTCAGATTATTCTGC
CCCCAATCCCATCAAAGTGACCCATCGGTGCTACATCCTTGAGAATGACACAGTCCAG
TGCGACTTGGACCTGTACAAGTCCCTGCAGGCTTGGAAGACCACAAGCTGCACATC
GACCATGAGATCGAAACCCTGCAGAACAAAATTAAGAACCTTCGAGAAGTCAGGGGT
CACCTGAAGAAGAAGCGACCGGAAGAATGTGACTGCCATAAAATCAGTTACCCACAGC
CAACACAAAGGCCGTCTCAAGCACAAAGGCTCCAGCCTGCACCCTTTCAGGAAGGG
TCTGCAGGAGAAGGACAAGGTGTGGCTGCTGCGGGAGCAGAAACGCAAGAAGAAA
CTGCGCAAGCTGCTCAAACGGCTGCAGAACAAACGATACGTGCAGCATGCCCGGCCT
CACGTGCTTTACCCACGACAACCACTGGCAGACGGCGCCACTCTGGACGCTGG
GGCCGTTCTGCGCCTGCACCAGCGCCAACAACAACACGTACTGGTGCTTGAGGACC
ATAAATGAGACCCACAACCTTCCTCTTCTGCGAATTTGCAACCGGCTTCATAGAATACTT
TGACCTCAGTACAGACCCCTACCAGCTGATGAACGCGGTGAACACACTGGACAGGG
ACGTCCTTAACCAACTGCACGTGCAGCTCATGGAGCTAAGGAGCTGTAAAGGCTACA
AGCAGTGCAACCCCCGGACCCGCAACATGGACCTGGGGCTTAGAGACGGAGGAAG
CTATGAACAATACAGGCAGTTTCAGCGTCGAAAATGGCCAGAAATGAAGAGACCTTCT
TCCAAATCACTGGGACAGCTATGGGAAGGTTGGGAAGGCTAAgcgccatagagagaggaac

FIG. 11A i

ctccaaaaccaggggcctcgtgtggctgccaggccatgcaaaaaacacccgattcccagaagatgaatgttgaact
gggagacctgacagaaggcagggcctgctcttgggacaggaaatcctggaggacagcgctggacttccgatgctca
gtttctttgccctgcttgcctggaatcaaacctcactggctgctctgggatgcgtgctcacacctggagtctctgctcacccttc
agaggctcaciaaagacaaaggaactaatttccatggacacttctccagagatggaaaattgctgggattcgcccactcct
cccctgcacccctccccagtcactaggaagcaagctgttttaaccttcttactcttggagaaagcacggacatcca
gggtgctgcaacctcacagtcttgacaaagtctatagcaciaaacagttaccattcaccaggctgggtgacctggctggctc
agaagctgccttcaccacatacatgaccgctcacacgtaaccaacacaggaattgtaggggaatctcactaatatgaa
atcccgctttcaagagtcgcggtgtcaataaacgctgtggctaggatcaaggataatcccttgagcttcagacatttattcct
gcccgggattcgttccttgttatccatcccagaactgatgttttctaaggtaccgaaaccccaagttgatgtgtcctgtgttt
aatgacattgtattgtaaagttttagtataagtaccatcttacagtggttctgccccagccaatgtctagctattggtatgaa
aaaaaaatctttgaattttgtaaaaaaaaaaaaaaaaa

FIG. 11A ii

FIG. 11A ii

>mouse SULF2 protein (translation of ORF)
MAPPGLPLWLLSTALLSLLAGSSAFLSHPRLKGRFQRDRRNIRPNILVLTDDQDVELG
SMQVMNKTRRIMEQGGAHFINAFVTPMCCPSRSSILTGKYVHNHNTYTNNECSS
PSWQAQHESRTFAVYLNSTGYRTAFFGKYLNEYNGSYVPPGWKEWVGLLKNSRFY
NYTLCRNGVKEKHGSDYSTDYLTDLITNDSVSFFRTSKKMYPHRPVLMVISHAAPHG
PEDSAPQYSRLFPNASQHITPSYNYAPNPDKHWIMRYTGPMKPIHMEFTNMLQRKR
LQTLMSVDDSMETIYDMLVETGELDNITYILYTADHGYHIGQFGLVKGKSMPYEFDIRV
PFYVRGPNVEAGSLNPHIVLNIDLAPTILDIAGLDIPADMDGKSILKLLDSERPVNRFHL
KKKLRVWRDSFLVERGKLLHKREGDKVNAQEENFLPKYQRVKDLCQRAEYQTACE
QLGQKWQCVEDASGTLKLHKCKGPMRFGGGGGSRALSNLVPKYDGQSSEACSCD
SGGGGDYKLGLAGRRKLFKKKYKTSYARNRSIRSVAIEVDGEIYHVGLDTPVQPRNL
SKPHWPGAPEDQDDKDGGSFSGTGGLPDYSAPNPIKVTHRCYILENDTVQCDLDLY
KSLQAWKDHKLHIDHEIETLQNKIKNLREVRGHLKKKRPEECDCHKISYHSQHKGRL
KHKGSSLHPFRKGLQEKKVWLLREQKRKKKLRKLLKRLQNNDTCSMPGLTCFTHD
NHHWQTAPLWTLGPFCACTSANNNTYWCLRTINETHNFLFCEFATGFIEYFDLSTDP
YQLMNAVNTLDRDVLNQLHVQLMELRSCKGYKQCNPRTRNMDLGLRDGGSYEQYR
QFQRRKWPEMKRPSSKSLGQLWEGWEG

FIG. 11B

100596 100596

FIG. 12

Genomic Organization of huSULF2 gene numbers represent base pairs

Contig	exon	start	end	length	gap
I 159532	1	66100	66668	568	28587
	2	95255	95529	274	20271
	3	115800	116039	239	34042
	4	150081	150232	151	

>13577

II 2152	5	1508	1677	169	
-------------------	---	------	------	-----	--

>5146

III 17546	6	4672	4822	150	1266
	7	6088	6263	175	4190
	8	10453	10581	128	1543
	9	12124	12180	56	455
	10	12635	12764	129	4101
	11	16865	17060	195	

>4971

IV 87036	12	4486	4714	228	308
	13	5022	5118	96	564
	14	5682	5776	94	1010
	15	6786	6845	59	508
	16	7353	7522	169	241
	17	7763	7905	142	1225
	18	9130	9253	123	2043
	19	11296	11329	33	245
	20	11574	11627	53	1007
	21	12634	13620	986	

FIG. 12